



## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/674,368  
Source: PCT/09  
Date Processed by STIC: 3/07/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

### Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:

<http://www.uspto.gov/web/offices/pac/checker>

# Raw Sequence Listing Error Summary

## ERROR DETECTED SUGGESTED CORRECTION

SERIAL NUMBER: 09/674,368

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 ☒ Wrapped Nucleics The number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 2 ☐ Wrapped Aminos The amino acid number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 3 ☐ Incorrect Line Length The rules require that a line not exceed 72 characters in length. This includes spaces.
- 4 ☐ Misaligned Amino Acid Numbering The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and use spacing between the numbers.
- 5 ☒ Non-ASCII This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.  
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
- 6 ☐ Variable Length Sequence(s) ☐ contain n's or Xaa's which represented more than one residue.  
As per the rules, each n or Xaa can only represent a single residue.  
Please present the maximum number of each residue having variable length and indicate in the (ix) feature section that some may be missing.
- 7 ☐ PatentIn ver. 2.0 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s) ☐. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies primarily to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 8 ☐ Skipped Sequences (OLD RULES) Sequence(s) ☐ missing. If intentional, please use the following format for each skipped sequence:  
(2) INFORMATION FOR SEQ ID NO:X:  
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")  
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:  
This sequence is intentionally skipped  
  
Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
- 9 ☐ Skipped Sequences (NEW RULES) Sequence(s) ☐ missing. If intentional, please use the following format for each skipped sequence.  
<210> sequence id number  
<400> sequence id number  
000
- 10 ☐ Use of n's or Xaa's (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.  
Use of <220> to <223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 11 ☐ Use of <213>Organism (NEW RULES) Sequence(s) ☐ are missing this mandatory field or its response.
- 12 ☐ Use of <220>Feature (NEW RULES) Sequence(s) ☐ are missing the <220>Feature and associated headings.  
Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"  
Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)
- 13 ☒ PatentIn ver. 2.0 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).  
Instead, please use "File Manager" or any other means to copy file to floppy disk.

PCT

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/674,368

DATE: 03/07/2001  
 TIME: 11:37:48

Input Set : A:\PTO.txt  
 Output Set: N:\CRF3\03072001\I674368.raw

Does Not Comply  
 Corrected Diskette Needed

5 <110> APPLICANT: American Cyanamid Company  
 7 <120> TITLE OF INVENTION: Vaccines Containing Recombinant Pilin Against Neisseria  
 9 Gonorrhoeae or Neisseria Meningitidis  
 11 <130> FILE REFERENCE: 33377-00/PCT  
 13 <140> CURRENT APPLICATION NUMBER: US/09/674,368  
 14 <141> CURRENT FILING DATE: 2000-10-27  
 16 <160> NUMBER OF SEQ ID NOS: 24  
 18 <170> SOFTWARE: PatentIn Ver. 2.0

*see item 5 on Error  
 Summary Sheet*

# ERRORED SEQUENCES

299 <210> SEQ ID NO: 13  
 300 <211> LENGTH: 31  
 301 <212> TYPE: DNA  
 303 <213> ORGANISM: Neisseria Meningitidis Class I  
 305 <400> SEQUENCE: 13  
 E--> 306 ccggcgcgtc tctcacggcg aatggccgg c  
 307 31  
 396 <210> SEQ ID NO: 24  
 397 <211> LENGTH: 22  
 398 <212> TYPE: PRT  
 399 <213> ORGANISM: Neisseria gonorrhoeae  
 404 <400> SEQUENCE: 24  
 405 Glu Ala Ile Leu Leu Ala Glu Gly Gln Lys Ser Ala Val Thr Glu Tyr  
 406 1 5 10 15  
 408 Tyr Leu Asn His Gly Lys  
 410 20  
 E--> 441 1  
 E--> 443 8  
 445 ACY-33377-00

*move up - format error  
 (see item 1 on  
 Error Summary  
 Sheet)*

*delete at end of file*

*see following pages for more errors*

BEST AVAILABLE COPY

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/674,368

DATE: 03/14/2001

TIME: 12:10:34

Input Set : A:\Pto.amc

Output Set: N:\CRF3\03142001\I674368.raw

2 <110> APPLICANT: American Cyanamid Company  
 4 <120> TITLE OF INVENTION: Vaccines Containing Recombinant Pilin Against Neisseria  
 6 Gonorrhoeae or Neisseria Meningitidis  
 8 <130> FILE REFERENCE: 33377-00/PCT  
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/09/674,368  
 C--> 11 <141> CURRENT FILING DATE: 2000-10-27  
 13 <160> NUMBER OF SEQ ID NOS: 24  
 15 <170> SOFTWARE: PatentIn Ver. 2.0  
 17 <210> SEQ ID NO: 1  
 18 <211> LENGTH: 504  
 19 <212> TYPE: DNA  
 20 <213> ORGANISM: Artificial Sequence  
 22 <220> FEATURE:  
 23 <223> OTHER INFORMATION: Description of Artificial Sequence:Chimeric of  
 24 Neisseria Meningitidis Class I and Neisseria  
 25 Gonorrhoeae  
 27 <220> FEATURE:  
 28 <221> NAME/KEY: CDS  
 29 <222> LOCATION: (1)..(501)  
 31 <400> SEQUENCE: 1  
 33 atg gat acc ctt caa aaa ggc ttt acc ctt atc gag ctg atg att gtg 48  
 34 Met Asp Thr Leu Gln Lys Gly Phe Thr Leu Ile Glu Leu Met Ile Val  
 35 1 5 10 15  
 37 atc gcc atc gtc ggc att ttg gcg gca gtc gcc ctt ccc gcc tac caa 96  
 38 Ile Ala Ile Val Gly Ile Leu Ala Ala Val Ala Leu Pro Ala Tyr Gln  
 39 20 25 30  
 41 gac tac acc gcc cgc gcg caa gtt tcc gaa gcc atc ctt ttg gcc gaa 144  
 42 Asp Tyr Thr Ala Arg Ala Gln Val Ser Glu Ala Ile Leu Leu Ala Glu  
 43 35 40 45  
 45 ggt caa aaa tca gcc gtt acc gag tat tac ctg aat cac ggc gaa tgg 192  
 46 Gly Gln Lys Ser Ala Val Thr Glu Tyr Tyr Leu Asn His Gly Glu Trp  
 47 50 55 60  
 49 ccc ggc aac aac act tct gcc ggc gtg gca tct tct tca aca atc aaa 240  
 50 Pro Gly Asn Asn Thr Ser Ala Gly Val Ala Ser Ser Ser Thr Ile Lys  
 51 65 70 75 80  
 55 ggc aaa tat gtt aag gaa gtt aca gtc gca aac ggc gtc att acc gcc 288  
 56 Gly Lys Tyr Val Lys Glu Val Thr Val Ala Asn Gly Val Ile Thr Ala  
 W--> 57 85 85 90 95 95 95 misaligned amino acid  
 59 aca atg ctt tca agc ggc gta aac aaa gaa atc caa ggc aaa aaa ctc 336  
 60 Thr Met Leu Ser Ser Gly Val Asn Lys Glu Ile Gln Gly Lys Lys Leu  
 W--> 61 100 105 110  
 63 tcc ctg tgg gcc aag cgt caa gac ggt tcg gta aaa tgg ttc tgc gga 384  
 64 Ser Leu Trp Ala Lys Arg Gln Asp Gly Ser Val Lys Trp Phe Cys Gly  
 W--> 65 115 120 125  
 67 cag ccg gtt acg cgc acc gac gcc aaa gcc gac acc gtc gcc gcc gcc 432  
 68 Gln Pro Val Thr Arg Thr Asp Ala Lys Ala Asp Thr Val Ala Ala Ala  
 W--> 69 130 135 140

BEST AVAILABLE COPY

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/674,368

DATE: 03/07/2001

TIME: 11:37:50

Input Set : A:\PTO.txt

Output Set: N:\CRF3\03072001\I674368.raw

L:13 M:270 C: Current Application Number differs, Replaced Application Number  
 L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
 L:60 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:64 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:68 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:72 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:76 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:80 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1  
 L:87 M:258 W: Mandatory Feature missing, <220> FEATURE:  
 L:87 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION:  
 L:194 M:258 W: Mandatory Feature missing, <220> FEATURE:  
 L:194 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION:  
 L:306 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:31 SEQ:13  
 L:441 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:24  
 M:332 Repeated in SeqNo=24

**BEST AVAILABLE COPY**

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/674,368

DATE: 03/14/2001  
TIME: 12:10:34

Input Set : A:\Pto.amc  
Output Set: N:\CRF3\03142001\I674368.raw

```

142   1               5               10               15
144 atc gcc atc gtc ggt atc ttg gca gcc gtc gcc ctg ccc gca tac caa   96
145 ile Ala ile Val Gly ile Leu Ala Ala Val Ala Leu Pro Ala Tyr Gln
146               20               25               30
148 gac tac acc gcg cgc gcc caa atg tcc gaa gcc ctg act ttg gca gaa   144
149 Asp Tyr Thr Ala Arg Ala Gln Met Ser Glu Ala Leu Thr Leu Ala Glu
150               35               40               45
152 ggt caa aaa tcc gca gtg atc gag tat tat tcc gac aac ggc aca ttc   192
153 Gly Gln Lys Ser Ala Val ile Glu Tyr Tyr Ser Asp Asn Gly Thr Phe
154               50               55               60
156 ccg aac agc aat act tcc gca ggt att gct gcc tct aac gag att aaa   240
157 Pro Asn Ser Asn Thr Ser Ala Gly ile Ala Ala Ser Asn Glu ile Lys
158   65               70               75               80
160 ggt aag tat gtg gca tgc gtt aag gtt gaa ggt aat gcc tct gtt gct   288
161 Gly Lys Tyr Val Ala Ser Val Lys Val Glu Gly Asn Ala Ser Val Ala
162               85               90               95
164 tct att acc gct acc atg aac tct agt aat gtg aat aag gac atc aaa   336
165 Ser ile Thr Ala Thr Met Asn Ser Ser Asn Val Asn Lys Asp ile Lys
166               100              105              110
169 ggt aaa acc ttg gta ctc gtc ggc aaa caa aac tcc ggt tgc gta aaa   384
170 Gly Lys Thr Leu Val Leu Val Gly Lys Gln Asn Ser Gly Ser Val Lys
171               115              120              125
173 tgg ttc tgc gga cag ccg gtt acg cgc gac aac gcc gac aac gac gac   432
174 Trp Phe Cys Gly Gln Pro Val Thr Arg Asp Asn Ala Asp Asn Asp Asp
175               130              135              140
177 gtc aaa gac gcc ggc aac aac ggc atc gaa acc aag cac ctg ccg tca   480
178 Val Lys Asp Ala Gly Asn Asn Gly ile Glu Thr Lys His Leu Pro Ser
179 145               150              155              160
181 acc tgc cgc gat acg tca tct gat gcc aaa   510
182 Thr Cys Arg Asp Thr Ser Ser Asp Ala Lys
183               165              170

```

186 <210> SEQ ID NO: 4  
187 <211> LENGTH: 170  
188 <212> TYPE: PRT  
189 <213> ORGANISM: Artificial Sequence

W--> 191 <220> FEATURE:  
W--> 191 <223> OTHER INFORMATION:  
191 <400> SEQUENCE: 4

```

192 Met Glu Ala ile Gln Lys Gly Phe Thr Leu ile Glu Leu Met ile Val
193   1               5               10               15
195 ile Ala ile Val Gly ile Leu Ala Ala Val Ala Leu Pro Ala Tyr Gln
196               20               25               30
198 Asp Tyr Thr Ala Arg Ala Gln Met Ser Glu Ala Leu Thr Leu Ala Glu
199               35               40               45
201 Gly Gln Lys Ser Ala Val ile Glu Tyr Tyr Ser Asp Asn Gly Thr Phe
202               50               55               60
204 Pro Asn Ser Asn Thr Ser Ala Gly ile Ala Ala Ser Asn Glu ile Lys
205   65               70               75               80
207 Gly Lys Tyr Val Ala Ser Val Lys Val Glu Gly Asn Ala Ser Val Ala

```

→ same env - see item 13  
on Env  
Summary Sheet

BEST AVAILABLE COPY

3

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/674,368

DATE: 03/14/2001

TIME: 12:10:34

Input Set : A:\Pto.amc

Output Set: N:\CRF3\03142001\I674368.raw

OK  
W--> 71 gcc aag acc gcc gac aac atc aac acc aag cac ctg ccg tca acc tgc 480  
72 Ala Lys Thr Ala Asp Asn Ile Asn Thr Lys His Leu Pro Ser Thr Cys  
W--> 73 145 150 155 160  
75 cgc gac gca agt gat gcc agc taa 504  
76 Arg Asp Ala Ser Asp Ala Ser  
W--> 77 165  
79 <210> SEQ ID NO: 2  
80 <211> LENGTH: 167  
81 <212> TYPE: PRT  
82 <213> ORGANISM: Artificial Sequence  
W--> 84 <220> FEATURE:  
W--> 84 <223> OTHER INFORMATION:  
84 <400> SEQUENCE: 2  
85 Met Asp Thr Leu Gln Lys Gly Phe Thr Leu Ile Glu Leu Met Ile Val  
86 1 5 10 15  
88 Ile Ala Ile Val Gly Ile Leu Ala Ala Val Ala Leu Pro Ala Tyr Gln  
89 20 25 30  
92 Asp Tyr Thr Ala Arg Ala Gln Val Ser Glu Ala Ile Leu Leu Ala Glu  
93 35 40 45  
96 Gly Gln Lys Ser Ala Val Thr Glu Tyr Tyr Leu Asn His Gly Glu Trp  
97 50 55 60  
100 Pro Gly Asn Asn Thr Ser Ala Gly Val Ala Ser Ser Ser Thr Ile Lys  
101 65 70 75 80  
104 Gly Lys Tyr Val Lys Glu Val Thr Val Ala Asn Gly Val Ile Thr Ala  
105 85 90 95  
108 Thr Met Leu Ser Ser Gly Val Asn Lys Glu Ile Gln Gly Lys Lys Leu  
109 100 105 110  
112 Ser Leu Trp Ala Lys Arg Gln Asp Gly Ser Val Lys Trp Phe Cys Gly  
113 115 120 125  
115 Gln Pro Val Thr Arg Thr Asp Ala Lys Ala Asp Thr Val Ala Ala Ala  
116 130 135 140  
118 Ala Lys Thr Ala Asp Asn Ile Asn Thr Lys His Leu Pro Ser Thr Cys  
119 145 150 155 160  
121 Arg Asp Ala Ser Asp Ala Ser  
122 165  
125 <210> SEQ ID NO: 3  
126 <211> LENGTH: 510  
127 <212> TYPE: DNA  
128 <213> ORGANISM: Artificial Sequence  
130 <220> FEATURE:  
131 <223> OTHER INFORMATION: Description of Artificial Sequence:Chimeric of  
132 Neisseria Meningitidis Class II and Neisseria  
133 Gonorrhoeae  
135 <220> FEATURE:  
136 <221> NAME/KEY: CDS  
137 <222> LOCATION: (1)..(510)  
139 <400> SEQUENCE: 3  
140 atg gaa gca atc caa aaa ggt ttc acc ctg atc gag ctg atg atc gtc 48  
141 Met Glu Ala Ile Gln Lys Gly Phe Thr Leu Ile Glu Leu Met Ile Val

see item 13 on  
Error  
Summary  
Sheet

BEST AVAILABLE COPY